



Extracting Densely Covered Areas Within Floating Car Datasets Using Inductive Loop Detector Data

Christian Röger¹ · Maja Kalinic¹ · Jukka M. Krisp¹

Received: 6 August 2020 / Accepted: 8 October 2020 / Published online: 5 December 2020
© The Author(s) 2020

Abstract

We present an approach to use static traffic count data to find relatively representative areas within Floating Car Data (FCD) datasets. We perform a case study within the state of Nordrhein-Westfalen, Germany using enviroCar FCD and traffic count data obtained from Inductive Loop Detectors (ILD). Findings indicate that our approach combining FCD and traffic count data is capable of assessing suitable subsets within FCD datasets that contain a relatively high ratio of FCD records and ILD readings. We face challenges concerning the correct choice of traffic count data, counting individual FCD trajectories and defining a threshold by which an area can be considered as representative.

Keywords Floating Car Data (FCD) · Inductive loop detectors (ILD) · Kernel density estimation (KDE) · Clustering

Extraktion von flächendeckend erfassten Gebieten in Floating Car Datensätzen (FCD) mit Daten induktiver Zähl Schleifen (ILD)

Zusammenfassung

Wir verwenden Daten aus Verkehrszählungen, um flächendeckend erfasste Gebiete aus Floating Car Datensätzen (FCD) zu extrahieren. Dafür führen wir eine Fallstudie in Nordrhein-Westfalen durch mit enviroCar FCD sowie Daten von induktiven Zähl Schleifen (ILD). Die Ergebnisse zeigen, dass durch unseren Ansatz Gebiete mit einem relativ großen Verhältnis von FCD und ILD Daten berechnet werden können. Es zeigt sich, dass die Wahl der benutzten Datensätze, der Zählprozess von FCD-Trajektorien an Zählerstandorten, sowie die Definition eines Grenzwertes, ab dem ein Gebiet als relativ flächendeckend erfasst gilt, relevant für die Aussagekraft der Ergebnisse sind.

Schlüsselworte Floating Car Data (FCD) · Induktive Verkehrszähl Schleifen (ILD), · Kernel Density Estimation (KDE) · Clustering

1 Introduction

A number of recently published reports show that individual car traffic in Germany has significantly increased since the 1990s.¹ According to several studies, traffic is affecting the climate on earth. Air pollution originating from the transportation sector contributes 25 percent to total Carbon Dioxide (CO₂) emissions (Mandl and Pinterits 2020) as well as 20%

to total Particulate Matter (PM_{2.5}) emissions.² This results in effects like global warming or health issues (Behera and Balasubramanian 2016). Other studies emphasize the formation of urban heat islands that emerge in weekly and daily periodical patterns, triggered by vehicular flow (Zhu et al. 2017).

Given the circumstances, data-driven intelligent transportation systems seek to utilize data resources efficiently

✉ Christian Röger
christian.roeger@geo.uni-augsburg.de

¹ University of Augsburg, Augsburg, Germany

¹ <https://www.deutschlandin zahlen.de/tab/deutschland/infrastinfrast/verkehr-und-transport/personenkilometer> (accessed September 28, 2020).

² <https://www.eea.europa.eu/data-and-maps/indicators/transport-emissions-of-air-pollutants-8/transport-emissions-of-air-pollutants-8> (accessed September 28, 2020).

and improve the performance and reliability of traffic services. Methods used to record traffic data have been evolving considerably and traffic information is increasingly becoming accessible worldwide.

Traffic count systems exist in different variations. In many cases, there are inductive loops detectors (ILD) installed on the pavement that sense electric changes when vehicles pass by. Traffic counters produce static data since they are installed on a specific point within a road network (Fabritiis et al. 2008).

Radio-frequency identification (RFID) transponders are among best methods to obtain individual travel times based on vehicle re-identification (Wright and Dahlgren 2001; Ban et al. 2010). In contrast to inductive loop detectors, these readers are located on the side of the road to keep record of the time the vehicle crosses a specific location. Similar to RFID, there exist also licence plate recognition systems (LPR) which are composed of cameras deployed along the roadway. These cameras identify licence plates using image-processing techniques and obtain the vehicle's travel time crossing two sensors. One such example is Oregon DOT's Frontier Travel Time project (Bertini et al. 2005).

Conventional methods of gathering data (such as static traffic counters like induction loops) have limits in obtaining near real time traffic information in large areas (Kalinic and Krisp 2019). All the above mentioned methods and techniques are costly to install and maintain, have limited coverage and provide only one (either density or travel time between two locations) of the three fundamental variables for traffic monitoring (density, flow and speed).

GPS devices (such as mobile phones) can, with high accuracy and frequency, obtain vehicle positions and instantaneous velocity readings. Since vehicles act as floating sensors, such acquisition method is called Floating Car Data. FCD, in contrast to static methods, provides traffic information for area-wide road networks. Assigning each individual vehicle and each single ride a specific ID, one can draw conclusions on the traffic situation.

Sanwal and Walrand (1995) and Zito et al. (1995) were one of the first to investigate and conclude that FCD constitutes a feasible source of data for traffic monitoring. Similar approach is seen in Moore et al. (2001), Schwarzenegger et al. (2008) and Bertini and Tantiyanugulchai (2004), who investigated the usage of dedicated vehicle fleets (such as FedEx and UPS trucks, taxis, busses) for traffic monitoring. This approach is particularly useful in larger cities.

Very often FCD readings are enriched with car diagnostic data such as speed, fuel consumption or torque. In that case, FCD is labelled as extended Floating Car Data (xFCD). xFCD data has been used in various environmental analysis (Fabritiis et al. 2008; Gühnemann et al. 2004; Ortenzi and Costagliola 2010; Röger et al. 2018) and included as a field

of study in the context of Location Based Services (LBSs; Huang et al. 2018).

FCD has emerged as a complement source of high quality data to the existing traffic detection technologies. Even though static counters are more precise when it comes to absolute vehicle counts, they are restricted in coverage and expensive to implement and maintain. FCD provides origin–destination trajectories for the entire road network. Currently the main drawback of FCD is their penetration rate, which evokes poor distribution and data density.

We present an approach for extracting relatively dense areas within Floating Car Datasets using records from static traffic counters. For assessing those kinds of coherent areas, we calculate the ratio between FCD readings and static counts. This generates a measure that gives us an indication of the quality for our FCD records. This paper includes a case study with Floating Car Data (FCD) and Inductive Loop Detector data (ILD). We apply common visual clustering methods like kernel density estimation (KDE) and compare them calculating the ratio between FCD trajectories and ILD records. In this paper, we combine extended Floating Car Data with traffic count data (obtained from Inductive Loop Detectors) for assessing relatively representative areas within the FCD dataset.

2 Analysis of Large Datasets

Nearly every digital process creates data. This includes readings from sensors, protocols and many more. Since digital storage such as server space or hard disks is considered affordable, it is convenient to store every accessible information possible. In many cases, this generates a huge and complex datasets that are challenging to analyse. Keim et al. (2006) even state that '[...] our ability to collect and store data is growing faster than our ability to analyse it'. This leads to a growing demand of methods for approaching and handling large-scale, multi-modal, and unstructured data (Keim et al. 2006).

We focus on a few, well-known and effectual clustering techniques, such as kernel density estimation (KDE) and spatial clustering approaches like *k*-means. All methods are applicable on variety of datasets; however, we use it to explore traffic data patterns.

2.1 Kernel Density Estimation (KDE)

Kernel Density Estimation (Silverman 1986) is widely used in visualizing and analyzing spatial data, with the objective of understanding event patterns (Smith et al. 2007). In KDE, series of estimations are made over a grid placed on the entire point pattern. They show the intensity at a certain location and therefore detect the highs and lows of event

pattern densities (Kalinic and Krisp 2018). Therefore, KDE is seen as a data-smoothing approach suitable for drawing conclusions about the distribution of population. Kernel Density operations approximate the probability density function f of a random variable X with n independent observations x_1, \dots, x_n . Therefore, the kernel density estimator $\hat{f}_h(x)$ at a point x is defined as:

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{j=1}^n K\left(\frac{x_j - x}{h}\right),$$

with K being the kernel function and h the bandwidth. Different kernel functions (e.g. Gaussian, Epanechnikov, Cosine, Triangular, to name a few) generate different results when performing KDE. Also, different bandwidth size plays a decisive role for the estimation. Therefore, it is necessary to experiment with these parameters to acquire map suitable for the user needs. There are several studies that utilize density estimation for region extraction within FCD. As an example, Tabibiazar and Basir (2011) detect congestion spots in FCD networks using a kernel-based density estimation.

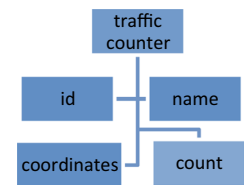
We use Kernel Density Estimation (KDE) method for finding dense areas within our datasets.

2.2 Clustering

Clustering is one approach to analyse spatial–temporal data at higher level of abstraction by grouping the data according to its similarity into meaningful clusters (Kisilevich et al. 2009). Overall, clustering methods are classified into four categories: partitioning method, hierarchical, density and grid based method. The common thing for all these methods is that the data is summarized by using local patterns, which describes the implicit relationship and characteristics of the data itself. The important part of clustering is to define the degree of similarity between movement data. One such example is seen in Liu and Ban (2013), which propose generating spatial clusters based on FCD trajectories at different time spans. These are further connected over time and space to form spatial–temporal clusters. Another study published by Yu et al. (2019) makes use of clustering for extracting urban functional regions out of FCD within the city of Chengdu, China.

To investigate performance of clustering methods on traffic data, we apply the commonly used k -means clustering method on our dataset. k -means is a great solution for pre-clustering, reducing the space into disjoint smaller subspaces where other clustering algorithms (if needed) can be applied. Additionally, it works really well with large datasets (Na et al. 2010).

Fig. 1 Data structure for traffic counters



3 A Case Study Using FCD and ILD

We will extract densely covered areas within a FCD dataset by visually exploring the data, performing a KDE, clustering the FCD trajectories and applying a novel approach by combining FCD and ILD. After extracting these densely covered areas, we compare the chosen techniques and approaches to find the best suitable for our dataset. We choose the federal state of Nordrhein-Westfalen located in western Germany as a research area. The chosen region has the highest count of inhabitants of all federal states within Germany.³ Additionally, Nordrhein-Westfalen has the highest population density of all federal states in Germany that are not a city-state.⁴ Consequently, we expect a good accessibility of data for this region. We use two major datasets that are described in the following.

3.1 Data

The first dataset, obtained from static traffic counters, is available at the German Federal Institute for Road Essence.⁵ The data is produced by automatic traffic counters. They are installed on motorways and federal highways. However, the dataset does not contain records obtained from traffic counters placed on inner-city roads. For Nordrhein-Westfalen region 281 traffic counters provide the access to the traffic data. Figure 1 illustrates the data structure for traffic counters.

Figure 1 shows that traffic counter data provides a unique ID and a name for each station as well as coordinates. Counts are separated for different vehicle types and not only contain information about the total number of cars. Therefore, separate counts of light and heavy vehicles are also accessible.

For the second dataset, Floating Car Data emerges from the enviroCar project (Bröring et al. 2015). It features a citizen science platform, where people can upload xFCD

³ <https://de.statista.com/statistik/daten/studie/71085/umfrage/verteilung-der-einwohnerzahl-nach-bundeslaendern/> (accessed September 28, 2020).

⁴ <https://de.statista.com/statistik/daten/studie/1242/umfrage/bevoelkerungsdichte-in-deutschland-nach-bundeslaendern/> (accessed September 28, 2020).

⁵ https://www.bast.de/BASt_2017/DE/Verkehrstechnik/Fachthemen/v2-verkehrszahlung/Daten/2016_1/Jawe2016.html?nn=1819490 (accessed September 28, 2020).

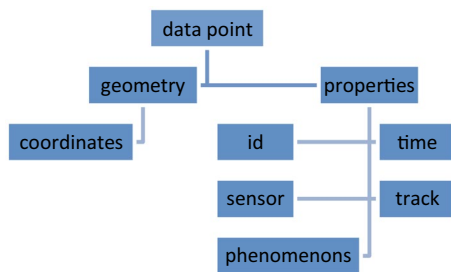


Fig. 2 Data point structure of the enviroCar dataset

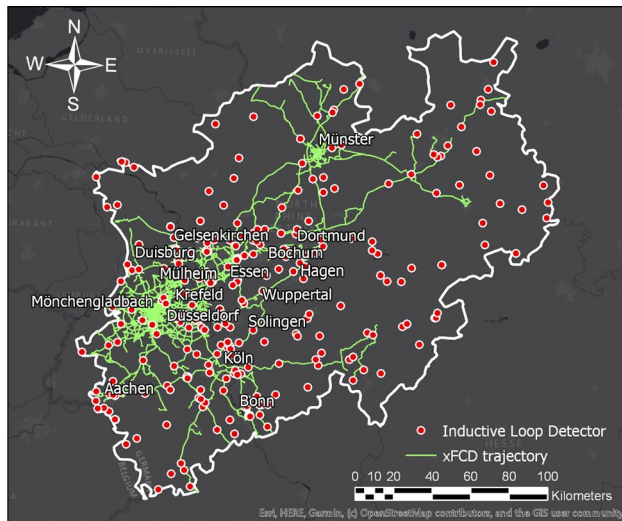


Fig. 3 Overview of the study area, Inductive Loop Counter (ILD) stations and extended Floating Car Data (xFCD)

produced by their own cars and download it using an API. We use data produced in 2016, since there are the most points accessible for the whole dataset. There are 2.934.018 FCD record points available. They emerge from 10.921 single trajectories. Figure 2 provides an overview for the enviroCar data structure.

Figure 2 shows that each record point comes with a certain structure represented by a JSON structure. At first, data points are divided into geometric information and properties. Latitude and longitude of a single measurement are stored within the geometry structure. The properties of a single measurement contain IDs for the measurement itself, an ID for the current trajectory, a timestamp as well as information on the driving car itself ('sensor') and readings from the On Board Diagnostics (OBD) interface ('phenomenons').

Figure 3 provides an overview of the two used datasets and the study area.

Figure 3 shows a map with the border of the study area, traffic counters illustrated by red circles ($n = 281$) and all available xFCD trajectories as green line features ($n = 10,921$, number of xFCD points = 2,934,018).

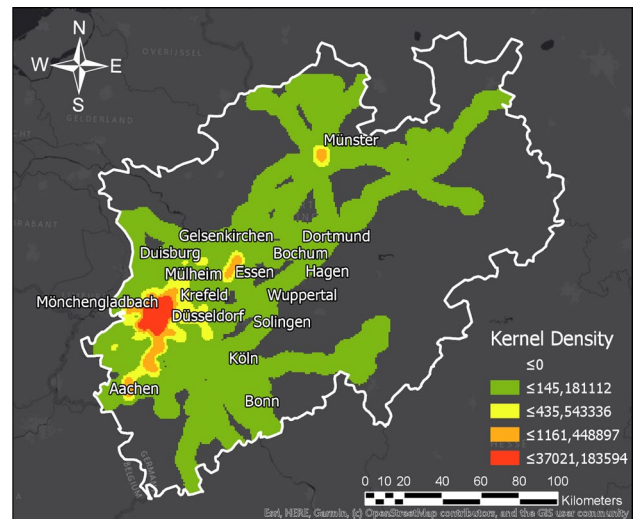


Fig. 4 Kernel density estimation (KDE) for our xFCD dataset

3.2 Gaining First Insights into the Data by Visually Exploring it

Visually exploring the FCD dataset may give hints where the distribution of data is sufficient and where not. Figure 3 indicates that there is an agglomeration of data points located in the western part of the study area. In addition, some highways leading southwards and in north-eastern direction from there provide data points.

Based on visual explorations we select the complete western and northern parts of the study area as a subset.

However, this kind of exploration only gives us an impression of the situation. We cannot see the exact amount of trajectories in each region. That is why we use a density-based approach for the next attempt.

3.3 Calculating Density Patterns Using KDE

Kernel density estimation (KDE) is used for estimating the probability density function of a variable. In our case, we apply KDE (with a grid size of 1000 m and a search radius of 460 m) for calculating the density of FCD within our study area. Figure 4 visualizes the results.

Figure 4 shows the kernel density of FCD record points with a green to red gradient. The densest areas emerge in the western and northern parts of the study area. It seems like most of the dataset origins from trajectories in the cities of Mönchengladbach, Essen and Münster as well as from highways.

Based on KDE, we choose the cities of Mönchengladbach, Essen and Münster as well as their connecting highways as a subset.

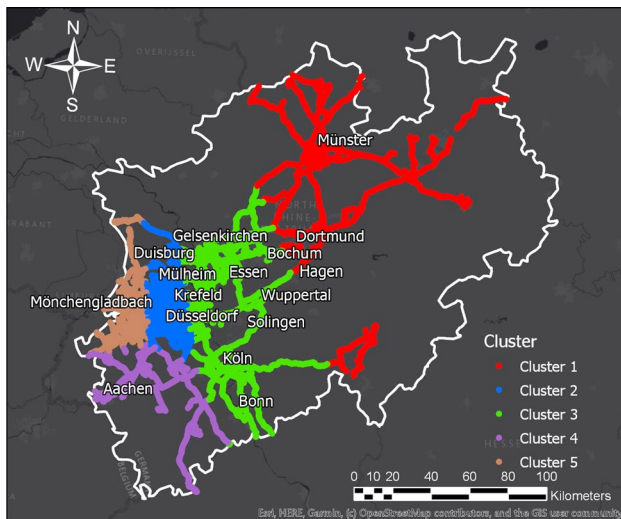


Fig. 5 *k*-Means clustering applied on our xFCD dataset

3.4 Spatially Clustering the Dataset Using *k*-Means

Cluster algorithms group data points into clusters based on the position of the features. As a well-known clustering technique, we apply *k*-means algorithm executed using KNIME for dividing our datasets into subsets. Figure 5 shows the results.

Figure 5 groups the dataset into five clusters. Consequently, five subsets of the dataset emerge. There are cluster one (red, 91,016 features), cluster two (blue, 2,136,741 features), cluster three (green, 156,527 features), cluster four (purple, 162,299 features) and cluster five (orange, 387,435 features). Cluster two and cluster five describe the densest clusters within our dataset.

Based on clustering, we select clusters number two and five as a subset of our dataset.

KDE and clustering show us the densest areas as well as coherent areas within our dataset. Yet they do not take into account the real traffic volume. The following approach therefore includes records from Inductive Loop Detectors.

3.5 Extracting FCD Clusters with Data from Inductive Loop Detector Data (ILD)

We would like to include the absolute numbers of cars passing by highways. Therefore, we count the number of individual FCD trajectories passing by all the traffic counters. We create a buffer around each counter station and count the number of trajectories within individual buffers. Next, we divide the calculated number of trajectories (at each counter station) with 365 since the static data shows daily counts of 2016 whereas FCD gives information on yearly traffic volume. Then, we calculate the ratio between mobile Floating

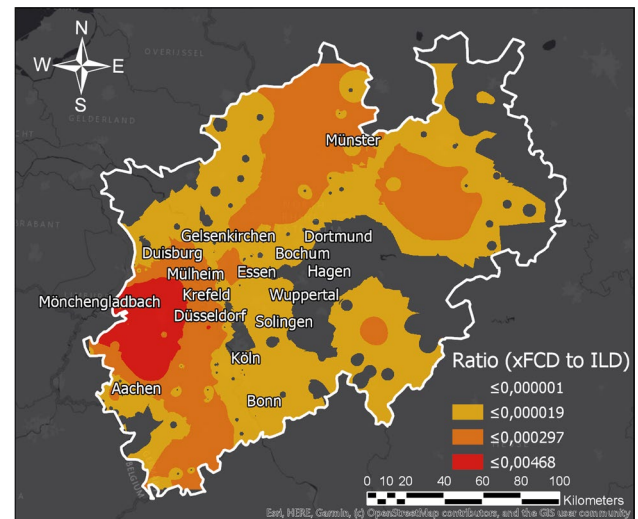


Fig. 6 Our novel approach using ILD counter readings applied on our xFCD dataset

Car Data and Inductive Loop Detector readings using following formula:

$$\text{ratio} = \left(\frac{\text{trajectories}_{\text{FCD}}}{\text{cars}_{\text{ILD}}} \right) \times 100.$$

As a result, we get a dataset with percentage information on how many FCD trajectories passed by counter stations relatively to the absolute daily count of vehicles. We interpolate our dataset using the inverse distance.

Weighted (IDW) tool. Based on this workflow, results are shown in Fig. 6.

Figure 6 shows the study area. There are four classes coded with a color range from orange to red. High values describe regions with a relatively high ratio of FCD records to ILD readings, and vice versa. This ratio is highest around the city of Mönchengladbach. Figure 7 provides a detailed view of the area with the highest calculated ratio.

Based on our approach we select the area around the city of Mönchengladbach as a subset of our dataset.

4 Results

Visual exploration reveals that the western and northern parts of our study area include more FCD points than their respective counterparts. To quantify this observation, we also applied a density analysis. It showed that most of the data points are located in the cities of Mönchengladbach, Essen and Münster as well as on the highways connecting those towns. A clustering analysis showed that areas in the western part of our study area around the cities of Mönchengladbach, Essen and Cologne provide a high

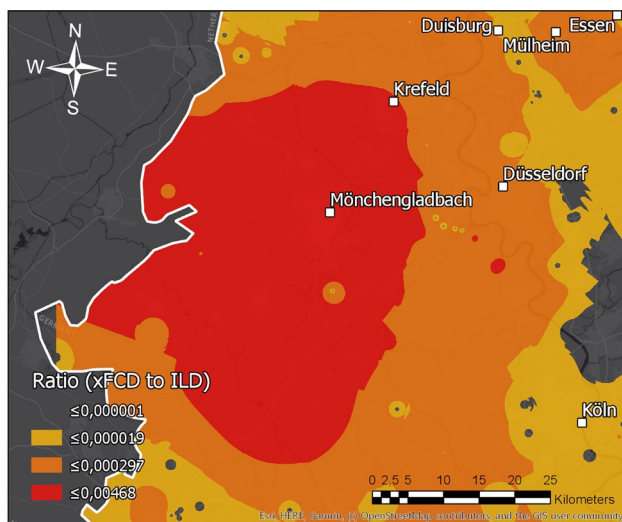


Fig. 7 Detailed view of the area around Mönchengladbach

agglomeration of data points. Our approach including the static traffic count data demonstrated that the highest ratio of FCD points to absolute vehicle counts is located in Mönchengladbach.

Visual exploration is suitable for a first insight into a dataset. It gives hints on where in the study area exist data and vice versa. It does not provide further quantitative information. Thus, we do not recommend visual exploration for choosing a study area.

Performing density based analysis (like KDE) creates better results. Densely covered areas get visible. In addition, this information is also quantifiable. Contrary to visual exploration, KDE showed that not the complete northern and western parts are covered entirely, as expected when using visual exploration. Also clustering provides good results for generating meaningful subsets of our dataset for performing further analysis. Still there is no consideration of real traffic numbers.

5 Discussion

The case study indicates that there are densely covered regions within the FCD dataset around the city of Mönchengladbach. Remke et al. (2018) substantiate this assumption, since there has been a project in Mönchengladbach encouraging citizens to contribute xFCD to the enviroCar database to optimize traffic light circuits in 2016.

The selection of suitable datasets matters. The traffic data has a dense distribution throughout the study area, yet traffic counters are only available for motorways and federal highways. Using a dataset that also features inner-city-bound data would have been better, since traffic also emerges on

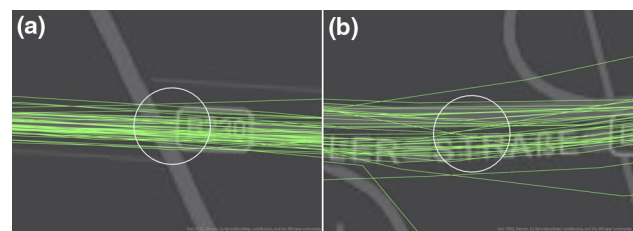


Fig. 8 Buffer that covers all FCD trajectories passing by (a) and buffer that does not cover every trajectory (b)

non-highway roads. Our approach calculates the ratio of FCD to ILD data by buffering the inductive loop detectors and counting the overlaying trajectories. Consequently, this technique is also applicable for ILD data obtained from inner-city traffic counters. The utilized Floating Car Data seems applicable, yet it does not cover the complete study area. In Fig. 5, there are some areas with no FCD at all.

Calculating the ratio between FCD trajectories and ILD readings provides us with useful results. In comparison to other studies that computationally extract regions out of FCD, as shown in Tabibiazar and Basir (2011) and Yu et al. (2019), we are not detecting areas with a certain traffic state or characteristic (e.g. traffic congestions or functional areas) but only focus on relatively dense covered regions. Analyses like congestion detection heavily rely on a high-density FCD dataset with many vehicles acquiring data. Our approach can filter those high-density subareas out of sparse FCD. However, there is a critical step when counting the number of FCD trajectories passing by traffic count stations. When buffering the traffic counter points, sometimes not every xFCD trajectory falls into the buffer polygons. Figure 8 provides us with an example for proper and problematic calculations respectively.

Figure 8 shows two examples with buffers (illustrated by white circles) used for calculating the number of FCD trajectories passing by ILD stations. Case (a) indicates a good performance since every FCD trajectory is covered by the buffer polygon. In case (b), the buffer is too small to cover all the lanes passing by the traffic counter. Using variable buffer diameters or map-matching FCD trajectories could solve the issue.

The visualization in Fig. 6 depends on the chosen threshold that indicates by which ratio (between FCD records and ILD counts) regions become relatively representable. When changing the symbology of the visualized data, regions that have not been visible before might get visualized as densely covered. Since datasets differ, it is a subjective choice which threshold to choose.

A scientifically correct conclusion which areas are absolutely densely covered and vice versa is not drawable.

6 Conclusions and Future Work

Our work concludes that there is a possibility of calculating densely covered areas within FCD datasets using static ILD data. As availability and size of datasets grows, data filtering gets more and more important. Our approach provides data analysts with a tool for filtering relatively dense areas out of big Floating Car Datasets. Consequently, we state that our work contributes to the field of analyzing traffic data. The choice of datasets and the technique to count numbers of FCD trajectories around counter stations (for calculating the number of FCD trajectories passing by) are critical. Also defining a threshold at which the ratio becomes representable and vice versa has been identified as challenging.

Future work will concentrate on improving the proposed workflow for calculating the ratio. There is a need to research further on how to count the number of FCD trajectories passing by traffic count stations. In addition, more datasets and study areas will be evaluated.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ban XJ, Li Y, Skabardonis A, Margulic J-D (2010) Performance evaluation of travel-time estimation methods for real-time traffic applications. *J Intell Transp Syst* 14(2):54–67
- Behera SN, Balasubramanian R (2016) The air quality influences of vehicular traffic emissions. In: *Air quality-measurement and modeling*, pp 113–133
- Bertini RL, Tantiyanugulchai S (2004) Transit buses as traffic probes: empirical evaluation using geo-location data. *Transp Res Rec J Transp Res Board* 1870:35–45
- Bertini RL, Lasky M, Monsere CM (2005) Validating predicted rural corridor travel times from an automated license plate recognition system: Oregon's frontier project. In: *Proceedings of the 2005 IEEE intelligent transportation systems*, 2005. IEEE, pp 296–301
- Bröring A, Remke A, Stasch C, Autermann C, Rieke M, Möllers J (2015) enviroCar: a citizen science platform for analyzing and mapping crowd-sourced car sensor data. *Trans GIS* 19(3):362–376
- de Fabritiis C, Ragona R, Valenti G (2008) Traffic estimation and prediction based on real time floating car data. In: *2008 11th international IEEE conference on intelligent transportation systems*. IEEE, pp 197–203
- Gühnemann A, Schäfer R-P, Thiessenhusen K-U, Wagner P (2004) Monitoring traffic and emissions by floating car data. Institute of Transport Studies, Australia
- Huang H, Gartner G, Krisp JM, Raubal M, van de Weghe N (2018) Location based services: ongoing evolution and research agenda. *J Locat Based Serv* 12(2):63–93
- Kalinic M, Krisp JM (2018) Kernel density estimation (KDE) vs. hot-spot analysis-detecting criminal hot spots in the city of San Francisco. In: *Proceeding of the 21st conference on geo-information science*
- Kalinic M, Krisp JM (2019) Fuzzy inference approach in traffic congestion detection. *Ann GIS* 25(4):329–336
- Keim DA, Mansmann F, Schneidewind J, Ziegler H (2006) Challenges in visual data analysis. In: *Tenth international conference on information visualisation (IV'06)*. IEEE, pp 9–16
- Kisilevich S, Mansmann F, Nanni M, Rinzivillo S (2009) Spatio-temporal clustering. In: *Data mining and knowledge discovery handbook*. Springer, pp 855–874
- Liu X, Ban Y (2013) Uncovering spatio-temporal cluster patterns using massive floating car data. *ISPRS Int J Geo-Inf* 2(2):371–384
- Mandl N, Pinterits M (2020) Annual European Union greenhouse gas inventory 1990–2018 and inventory report 2020. Edited by European Environment Agency. Available online at <https://www.eea.europa.eu/publications/european-union-greenhouse-gas-inventory-2020>. checked on 9/29/2020
- Moore I, James E, Cho S, Basu A, Mezger DB (2001) Use of Los Angeles freeway service patrol vehicles as probe vehicles. California PATH, Richmond
- Na S, Xumin L, Yong G (2010) Research on k-means clustering algorithm: an improved k-means clustering algorithm. In: *2010 third international symposium on intelligent information technology and security informatics*. IEEE, pp 63–67
- Ortenzi F, Costagliola MA (2010): a new method to calculate instantaneous vehicle emissions using OBD data. SAE technical paper
- Remke A, Rieke M, Stasch C, Autermann C, de Wall A, Wulffius H, Jirka S (2018) Open traffic analysis. In: *EGUGA*, p 19448
- Röger C, Keler A, Krisp JM (2018) Examining the influence of road slope on carbon dioxide emission using extended floating car data. In: *Adjunct Proceedings of the 14th international conference on location based services*. ETH Zurich, pp 135–140
- Sanwal K, Walrand J (1995) Vehicles as probes (No. California Path Working Paper UCBITS-PWP-95-11). University of California Berkeley, Institute of Transportation Studies, Berkeley, CA
- Schwarzenegger A, Bonner DE, Iwasaki RH, Copp R (2008) State highway congestion monitoring program (HICOMP). In: *Annual data compilation*, Sacramento, CA
- Silverman BW (1986) Density estimation for statistics and data analysis. CRC Press, Boca Raton, p 26
- de Smith MJ, Goodchild MF, Longley P (2007) Geospatial analysis: a comprehensive guide to principles, techniques and software tools. Troubador Publishing Ltd, Kibworth Harcourt
- Tabibiazar A, Basir O (2011) Kernel-based modeling and optimization for density estimation in transportation systems using floating car data. In: *2011 14th international IEEE conference on intelligent transportation systems (ITSC)*. IEEE, pp 576–581
- Wright J, Dahlgren J (2001) Using vehicles equipped with toll tags as probes for providing travel times. California PATH Working Paper
- Yu B, Wang Z, Mu H, Sun Li, Hu F (2019) Identification of urban functional regions based on floating car track data and POI data. *Sustainability* 11(23):6541
- Zhu R, Wong MS, Guilbert É, Chan P-W (2017) Understanding heat patterns produced by vehicular flows in urban areas. *Sci Rep* 7(1):1–14
- Zito R, D'este G, Taylor MAP (1995) Global positioning systems in the time domain: how useful a tool for intelligent vehicle-highway systems? *Transp Res Part C Emerg Technol* 3(4):193–209